

Frames Program Overview

Michael Ratcliffe, Senior Advisor for Frames, Geography Division

March 19, 2021

Shape
your future
START HERE >

United States®
Census
2020

Frames Program

The Challenge

- Various frame-like datasets are in use at the Census Bureau.
- Contain information about addresses, businesses, jobs, and people.
- Some information is unique within a particular frame.
- Other information is duplicative (e.g., group quarters exist in both the Master Address File and Business Register).
- Frames exist in an uncoordinated and unintegrated environment.
- No process exists that allows for the direct linkage of information contained in one frame with information in any other frame.

Frames Program

The Vision

To create Enterprise-wide frames linkable in nature, agile in structure, accessible for production or research on a need-to-know basis, and that adhere to best practices in terms of technology usage, data management, and methodology.

Linkable in Nature

- Each frame will include the necessary unique identifiers and keys for linkage to each other. For example:
 - Location information will be available on each person, job, and business record.
 - A person record may have linkage information for a job.
 - A job record may have linkage information to a business.

Frames Program

Benefits of the Vision

- Fosters an environment to do more work in the office.
- Reduce burden on respondents (individuals and businesses) by re-using data that already exist in one or more frames.
- Reduce duplication, manual efforts, and one-off data cleaning projects.
- Allows more time for production and research.
- Allows for ability to more easily trace data and code through the linked frames, facilitating reproducibility.
- Expands the Census Bureau's ability to report on changes and trends affecting the Nation's population, economy, and communities.

Frames Program

Master Address File (MAF) – Current State

- National address list of housing units, group quarters, transitory locations and some non-residential addresses.
- Includes over 200 million address records.
- Linked to the Topologically Integrated Geographic Encoding and Referencing (TIGER) System.
- Serves as the source of address and location information for the Decennial Census, American Community Survey, and other demographic surveys.
- Updated using information from the U.S. Postal Service; tribal, state, and local governments; and Census data collection operations.

Frames Program

Business Register (BR) – Current State

- National inventory of businesses.
- Includes over 30 million businesses.
- Serves as the source of sampling and publication for the Economic Census.
- Updated using data from the Internal Revenue Service (IRS) tax records; the Bureau of Labor Statistics; the Social Security Administration; the Economic Census; and some Census surveys.
- Updated using geographic information from the MAF/TIGER System.

Frames Program

Job Frame – Current State

- An inventory of “job records.”
- Each “job record” is a relationship between a person and a firm in which at least \$1 of wages were paid from the firm to the worker in a given quarter.
- Over 14 billion job records.
- Covers over 97 percent of private employment and most federal, state, and local jobs.
- Constructed and maintained for public-use data products.
- Additionally, the LEHD Program maintains a business frame that is independent of the BR, known as the Employer Characteristics File (ECF).

Frames Program

Demographic Frame – Current State

- Several disparate person frames exist within the Census Bureau.
- 2010 Match Study evaluated coverage and quality of administrative data relative to the 2010 Census.
- 2020 Census uses administrative records to support 16 operations.
- The 2020 Administrative Records Census (part of the 2020 Census evaluation program) will be a de-duplicated count of all individuals found in administrative records on Census Day.
- LEHD Individual Characteristics File contains a record for every person ever employed in an LEHD state, participated in the LEHD program, and captured in the state's unemployment insurance records.

Frames Program

Use Cases

- Update and maintenance of address use type (residential or non-residential or both) on the MAF based on information in the Business Register, Job Frame, and Demographic Frame.
- Identification of group quarters using information from the MAF, Business Register, and the Job Frame.
- Reporting on demographic and economic impacts of federal programs, such as the Opportunity Zone and Enterprise Zone Programs, and the extent to which programs are achieving their goals.
- Reporting on demographic characteristics of business owners and their employees in types of establishments within a specific geographic area (that is, more specific than industry or occupation).

Frames Program

Organizational Structure

Program Management Office

- Michael Ratcliffe, Senior Advisor
- Lori Zehr, Senior Technical Lead
- Kayla Curcio, Program Manager

Frames Governance Group

- Ron Jarmin, Co-Chair
- Deirdre Bishop, Co-Chair

Subject Matter and IT Teams

- Geospatial Frame Team – Richard Watson
- Business Frame Team – Jessica Wellwood
- Job Frame Team – Matthew Graham
- Demographic Frame Team – Jennifer Ortman
- Chief Technology Office– Pura Perez and Chris Carrino

Frames Program

FY20 Accomplishments: Laying the Foundation

- Laying the project management foundation in advance of official program status:
 - Developed a roadmap for the future.
 - Prepared “As-Is” documentation for each of the four frames as well as a high-level overview.
 - Drafted “To-Be” documentation for overall Enterprise Frames and for each of the four frames.
 - Drafted vision canvasses, epics, and identified use cases.
 - Identified potential linking keys and scoping potential requirements for linking frames.
 - Identified FY21 and FY22 Annual Performance Plan Indicators and Milestones.
- Engaging with internal and external stakeholders/potential users of linked frames.
- Collaborating with, and participating on, various enterprise teams and efforts.
- Identified the organizational structure for the Frames Program as well as staffing plans for each team within the program.

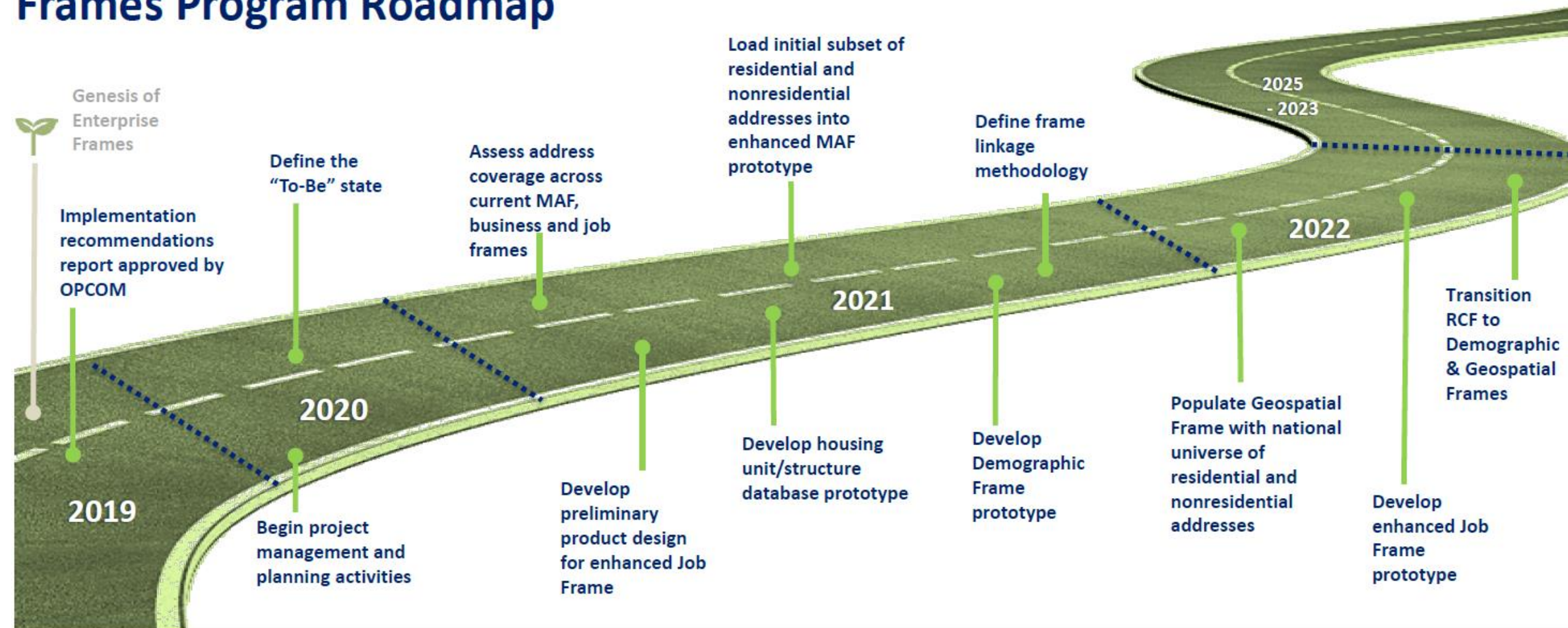
Frames Program

Stakeholder Engagement and Discussion

Meet with stakeholders to discuss

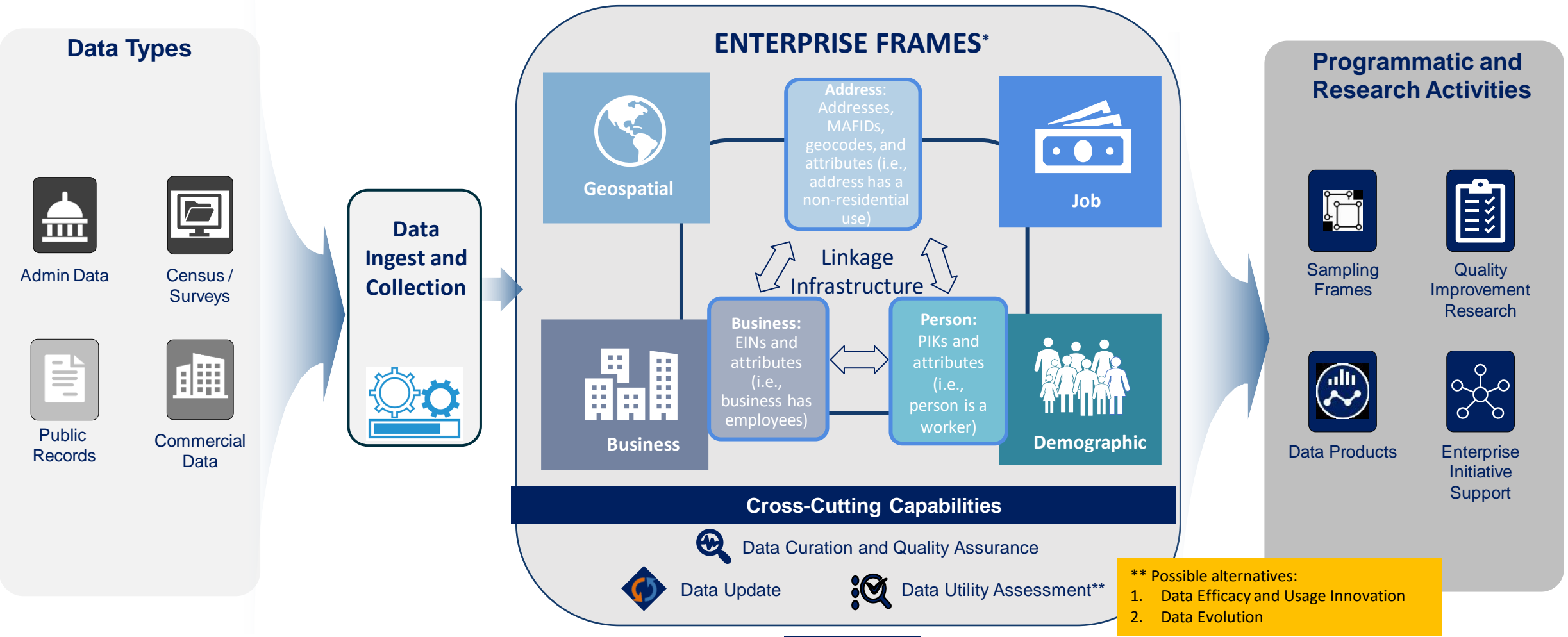
- Ways in which linked frames can improve survey efficiency, sampling, and potentially reduce respondent burden.
- Ways in which to reduce redundancies and duplication of effort regarding internal Census Bureau access and use of data provided by federal agency partners.
- Feasibility of carrying out Frames Program research and development tasks under existing approved projects.
- Potential for use agreements providing more flexible use of demographic and economic data to update and maintain frames, for internal research and development purposes, and for potential use in household and economic surveys to fill in missing data or corroborating responses.

Frames Program Roadmap



Program Goal: The vision of the Frames Program is to create enterprise-wide frames linkable in nature, agile in structure, accessible for production or research on a need-to-know basis and that adhere to best practices in terms of technology usage, data management, and methodology.

Creating an Infrastructure to Modernize the Census Bureau's Statistical Foundation



- Assumptions**

 - Data agreements are in place to support enterprise data sharing
 - Data management practices are well established and followed
 - Policy and access rules are enforced
 - Best practices/methods are well-documented and are available for review and re-use
 - Evaluation and improvement of data and methodologies is performed continuously

- Dedicated research staff is available
 - Data is available upon ingestion
 - Ingest once, allow access across frames
 - Built in “real-time-updates” triggers across the frames
 - IT issues resolved for efficiency and security and usability

*This diagram provides a generalized view of the two-way flows of specified data elements (MAFIDs, PIKs, EINs, addresses, geocodes, etc.) between pairs of frames.

Frames Program

Questions

1. We currently are considering using Person Identification Keys (PIKs), Employer Identification Numbers (EINs), and Master Address File Identifiers (MAFIDs) as the primary linking keys between frames. Are there other data elements that we should consider using as linking keys? Are there other methods for linking across the frames we should consider?
2. What thoughts and advice do Committee members have regarding methods for measuring the quality and completeness of each frame? What measures can we put in place to determine whether a frame is comprehensive and as-complete-as possible?
3. How do we build efficient geospatial relationships into the linking process; that is, efficient methods for locating all records within a specified higher-level geographic area or areas, more effectively linking records in the MAF to geographic areas in the TIGER database?

Thank you!